Challenges for Enterprise GIS

Gordon N. Keating, Paul M. Rich, Marc S. Witkowski

GISLab, Environmental Dynamics & Spatial Analysis Group, EES-10
MS-D452, Los Alamos National Laboratory, Los Alamos, NM 87545 USA
tel: 505-667-5902, fax: 505-667-1628, e-mail: gkea@lanl.gov

LANL Technical Information Release Number LA-UR-02-1830

ABSTRACT

A natural result of the widespread growth in geospatial data use in large organizations is the development of institution-wide, or "enterprise," geographical information systems (GIS). A key issue in this development is effective geospatial data sharing within and across organizational boundaries. In the aftermath of the destructive May 2000 Cerro Grande wildfire, new ideas for enterprise GIS and geospatial information management are being considered at Los Alamos National Laboratory (LANL). Barriers to change and challenges to implementation of enterprise GIS make the transition difficult. These challenges range from organizational divisions to budgets and data standards. The purpose of this paper is to offer observations and analysis based on our experience in designing a GIS for the Cerro Grande Rehabilitation Project (CGRP-GIS). As we examine these challenges and the solutions being developed for this project, the foundation is laid for developing a LANL-wide, enterprise GIS design, offering benefits for research, emergency management, operations, and natural hazards mitigation. The CGRP-GIS serves as a prototype for enterprise GIS design.

INTRODUCTION

Despite rapid growth in the use of geographical information system (GIS) technologies in public and private organizations, recent advances in data storage, processing, and networking do not necessarily result in increased data accessibility. With the goal of enhanced geospatial data sharing within and across organizational boundaries, organizations increasingly focus on enterprise, or institutional, solutions to effective information exchange (Meredith 1995, Burrough and McDonnell 1998), thereby avoiding redundant systems and services and incompatible infrastructure (Pinto and Onsrud 1995). Enterprise GIS is a virtual rallying call, and this paper offers evidence that it is the preferred and most effective solution to these needs. The key to successful enterprise GIS is development of an appropriate organizational framework (Oppman 1999). In this paper we focus primarily on data sharing issues, though enterprise GIS also includes shared information and analysis resources (Davis, 1999).

The evolution of data management in large organizations typically follows a "punctuated equilibrium" model (Gould and Eldredge 1977), in which the status quo limits growth or change until the system is disturbed and then rapid change occurs, followed by a new status quo. Such a disturbance may come in the form of a natural or man-made disaster, during which urgency demands action, the limitations of the existing system are exposed, and necessity opens the door for change. In the aftermath of the destructive May 2000 Cerro Grande wildfire (Salazar-Langley et al. 2000a, Salazar-Langley et al. 2000b, Webb and Carpenter 2001), GIS and information management (IM) experts at Los Alamos National Laboratory (LANL) are considering options for improved geospatial data management and information exchange, based on lessons learned during and shortly after the fire. Herein we examine challenges faced when building enterprise GIS at LANL, as well as potential solutions. The experiences of the GIS community at LANL illustrate the difficulties and opportunities of any large institution during a time of transition.

The Cerro Grande Rehabilitation Project (CGRP) was launched at LANL shortly after the May 2000 Cerro Grande wildfire. The CGRP goals are to restore infrastructure, to inform the public, to assist scientists and emergency managers in assessing long-term environmental impacts, and to mitigate future hazards associated with the aftermath of the wildfire. As a result of the firefighting and subsequent rehabilitation efforts, massive amounts of geospatial data are being generated. These data detail the extent, severity, and progression of the fire; condition of soils, vegetation, archaeological sites, and LANL facilities; and management treatments for slope stabilization, flood mitigation, and revegetation. Geospatial data are being generated to support decisions about post-fire recovery, to mitigate floods and other hazards, and to document potential environmental impacts.

A GIS component of the CGRP was charged with capturing and managing geospatial data associated with the fire fighting, rehabilitation, and hazard mitigation efforts; to provide rapid access to and visualization of the data; and to integrate the data into predictive models and risk assessment systems. This CGRP-GIS effort focuses on building information management infrastructure, including a central geospatial data repository and an Internet-based mapping interface, along with data management policies and procedures to ensure data quality and streamline access by many users at LANL. The effort to build a GIS to serve the CGRP has highlighted the overarching challenges to building any enterprise GIS. Challenges range from building the necessary infrastructure, in terms of hardware, software, and comprehensive databases, to the human element, in terms of collaboration and data sharing.

We begin with a summary of events during and following the Cerro Grande Fire and then examine common needs and goals of the major stakeholders of the CGRP-GIS. Within the framework of an idealized cycle of geospatial data management, we examine the need for enterprise GIS, challenges for design and implementation, and a prototype solution. The special needs of the Emergency Operations Center (EOC) are then considered, both in the broad conception of the emergency management industry and in the particular case of the new LANL EOC, currently in the design phase. Finally, we provide a synthesis of our findings and offer conclusions.

BACKGROUND

The Cerro Grande Fire

The Cerro Grande Fire began as a prescribed burn on Thursday, May 4, 2000, set by Bandelier National Monument personnel to reduce brush and reestablish native vegetation (Figure 1). That night the fire slipped out of control and became a wildfire, heading north and east from its starting point on Cerro Grande (Mountain), southwest of Los Alamos. By Sunday, May 7, 50-mile-per-hour winds whipped the fire out of control, the LANL EOC was activated, and the Western Area neighborhood of Los Alamos was evacuated. By that evening, LANL administration announced emergency closure for Monday, May 8, and LANL employees did not resume normal work schedules until Monday, May 22. On Wednesday, May 10, conditions grew more dangerous, and the entire town of Los Alamos was evacuated as 50- to 70-mph winds drove the fire across fire lines and into town, eventually burning 15,500 acres in 9 hours. Later that night, the town of White Rock was also evacuated, bringing the total number of displaced people to 18,000.

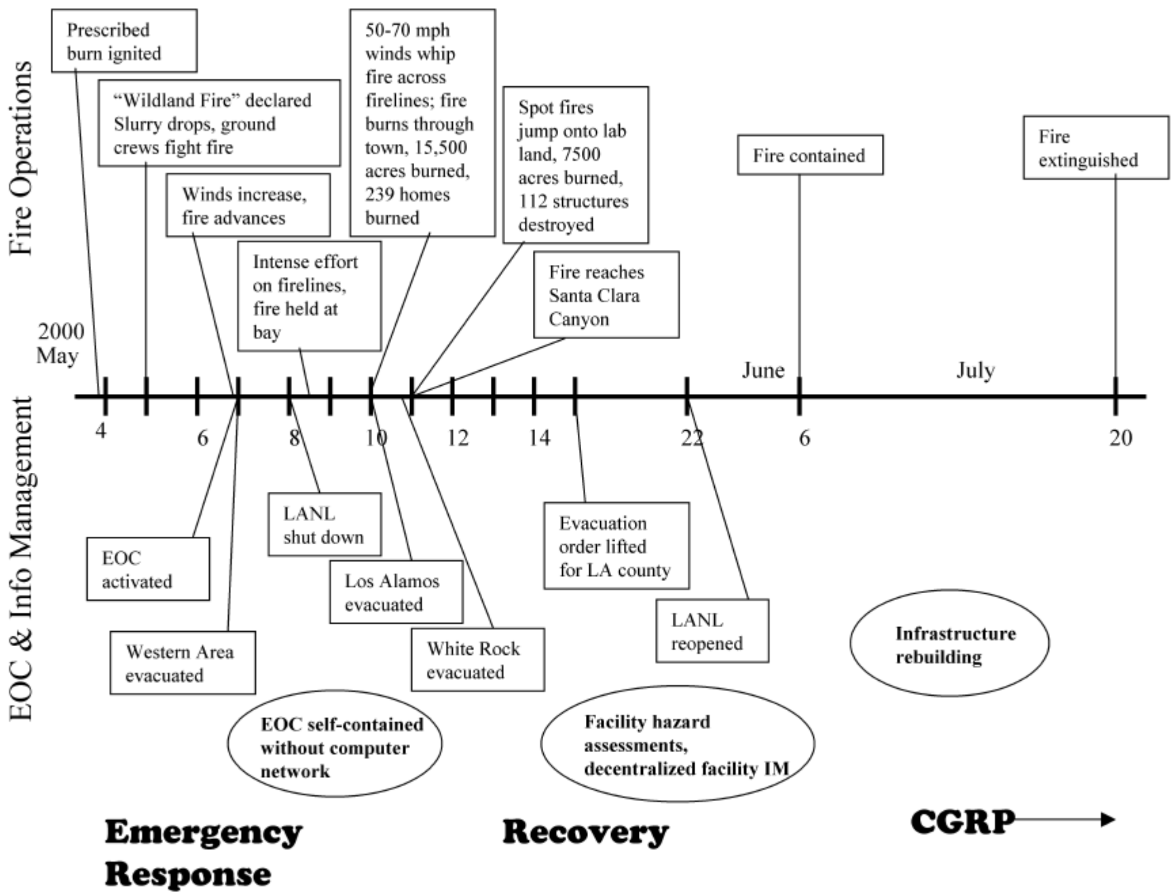

Figure 1. Timeline for Cerro Grande fire, highlighting fire operations activities (top) and the status of the LANL Emergency Operations Center (EOC) and information management (bottom) (after Webb and Carpenter 2001).

The fire was officially contained on June 6, a month after it started, though it was not extinguished until July 20. During this extraordinary period, over 48,000 acres of Federal, County, Tribal, and private land burned, including 8,000 acres of LANL land and 112 LANL structures. In Los Alamos, 239 buildings were destroyed, displacing 429 families. Total losses and expenses associated with the fire-fighting and recovery efforts exceeded $1 billion (Salazar-Langley et al. 2000a, Webb and Carpenter 2001).

During the fire, the EOC was a hub for communications, information, and emergency management. Emergency managers required current data on facilities, infrastructure, LANL hazard areas, terrain, meteorological conditions, fire status, and whereabouts of workers in the field. Leaders from many areas of the lab converged on the EOC to bring expertise and aid in crisis decisions. An emergency GIS operations center was set up in the nearby city of Santa Fe and was staffed with LANL and Sandia National Laboratory personnel (Mynard et al. 2002). Unreliable electrical power and computer networks made external information hard to access, and the majority of decisions had to be made based on information stored at the EOC and recovered by GIS experts, even though more current information existed elsewhere at LANL.

Further challenges were posed by incompatible communications systems among the many agencies involved in the fire-fighting effort.

As the active front of the fire passed to the north and east, away from the main LANL technical areas, the LANL firefighting effort changed into recovery work. The EOC continued to be the nexus for information and coordination. The main focus of the recovery efforts at LANL was the safe restart of normal operations. This included hazard assessments of all facilities and formal approval for occupation and work in LANL buildings. Because LANL had never before been confronted with the complete evacuation of the site (and county) or an emergency of this magnitude, the lack of a comprehensive emergency plan hampered the efficiency of the restart process (Salazar-Langley et al. 2000a). No central priority list existed to guide the efforts at reopening LANL buildings. The decentralized nature of facilities management, with its distributed and redundant information management systems, led to concerns about data accuracy, access, and consistency (Salazar-Langley et al. 2000b).

The "recovery" effort evolved into the Cerro Grande Rehabilitation Project (CGRP) in late June 2000. With the recognized limitations in availability and quality of geospatial data during the fire crisis and recovery periods, an important component to the multi-million-dollar CGRP effort is the development of an efficient GIS and information management system for the CGRP and the new EOC. This project, the CGRP-GIS, is tasked with building a geospatial data warehouse for fire-related projects, including efficient data storage and access, reliable data quality, and adherence to national metadata standards. Given the challenges posed by inadequate institutional GIS resources during and shortly after the Cerro Grande fire, an enterprise GIS at LANL could be viewed as a phoenix rising from the ashes.

Existing GIS Structure at LANL

Before the Cerro Grande fire, GIS at LANL consisted of project-level systems with limited, *ad hoc* sharing of data and other resources. Spatial information management practices were different in separate GIS for planning, utilities and infrastructure, emergency management, environmental monitoring and restoration, and earth science research. Since the CGRP was by definition an institution-wide effort, the design of the CGRP-GIS began with an assessment of the spatial data and information needs of the institution for fire rehabilitation. An evaluation of existing resources followed that highlighted the diverse needs of the various GIS stakeholders at LANL.

The form of the CGRP-GIS and the potential LANL enterprise GIS depends on the composition of its "stakeholders," those who have a share or interest in geospatial data at LANL. GIS stakeholders fall into four general categories: data providers, data managers, GIS users, and customers (Figure 2). Data providers are those who generate new geospatial data and serve as the steward or owner, responsible for these particular data. Data managers maintain spatial data warehouses and need consistent workflow procedures that ensure efficient and standardized means to manage and deliver data. Data users are GIS professionals and analysts who need access to many kinds of geospatial data, recombine data in the course of their work, and often generate new, updated or derived, datasets. Customers pay for particular GIS products or sets of work. GIS stakeholders share common needs, such as data quality standards, data documentation ("metadata"), consistent data formats, and data archiving. However, each stakeholder has unique goals and requirements, in terms of infrastructure (hardware/software) and data, as well as unique preconceptions about the value of enterprise GIS to his or her goals and willingness to participate in data sharing.

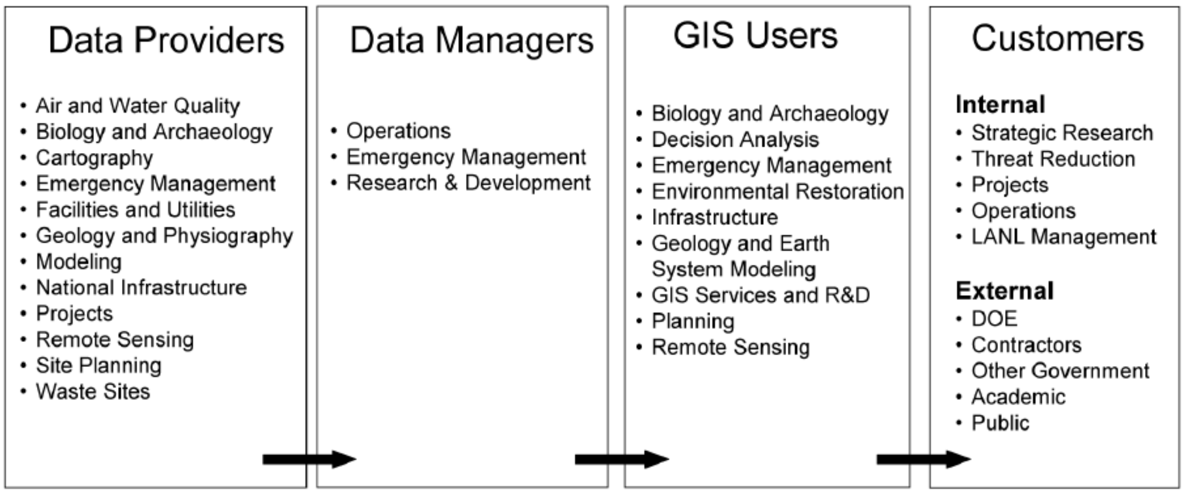

Figure 2. GIS stakeholder roles and functional differences at LANL.

Specific stakeholders may play a single role or multiple roles in enterprise GIS (Figure 2).

LANL Management. LANL management is responsible for institutional policies, oversees work performed, and makes funding decisions. Until recently GIS received little attention or recognition by top-level managers and has mostly been supported as relevant to the LANL mission at the project level. Top-level management has a key role in the eventual success or failure of any enterprise GIS. During crises, management supports the emergency activities by handling public information regarding operations status and potential environmental impacts during and after the emergency. The GIS and geospatial data needs of LANL management are mostly derivative, in the form of status reports and summary graphics, including maps and tabular analyses, and as such management is a customer. Management depends on data providers and data users from all parts of LANL, including operations and research GIS professionals.

Emergency Managers. Emergency managers coordinate response to natural and anthropogenic emergencies at LANL, serving as the central nervous system of crisis operations. Their responsibilities include management of environmental, safety, and security issues. For larger crises, this group maintains the EOC, where representatives from LANL management and operations converge to coordinate activities during an emergency.

Because of the uncertain and potentially wide-ranging impact of natural and human-induced hazards, emergency managers at the EOC require efficient access to diverse types of geospatial and tabular data at LANL, and therefore they have comprehensive GIS and IM needs and play the role of data managers, GIS users, and customers. These data types include topographic and environmental information; roads, buildings, utilities, electrical, and security layers; and tabular information regarding facility use, potentially hazardous substances, and human resources. Results of predictive simulations (e.g., flooding, atmospheric dispersal, wildfire, etc.) may also be helpful during crises. While the EOC managers cannot anticipate all data needs, each member of the crisis command team (from diverse areas of LANL management) understands key supporting information he or she may require to make informed decisions during a crisis. Each can provide guidance to help develop a complete data repository.

As was demonstrated during the Cerro Grande fire, the EOC must be self-contained during emergencies and cannot depend on access to distributed data sources. Therefore, system requirements at the EOC include large-volume data storage devices, database and software servers, networked workstations, visualization and cartographic software, and efficient off-site backup of critical data.

Operations. Operations personnel maintain infrastructure, handle raw materials and waste, coordinate fire protection, and plan and implement improvements and new construction at LANL. During crises, facility managers team up with security and other emergency managers at the EOC to make informed decisions and coordinate emergency activities.

Geospatial data needs for operations are centered on infrastructure: roads, utilities, structures, hazardous materials etc.; however, physiographic and environmental data are important as well (e.g., topography, vegetation, geology, soils, waste sites, plant and animal habitat, archaeological sites, erosion potential, and runoff). Supporting tabular data include information on projects underway in various facilities, personnel, hazardous material lists, and contact lists. Operations personnel generate much of their own data, but they also draw heavily on data and expertise from throughout LANL (orthophotos, geology, topography, etc.). The utilities mapping team provides facilities location data and as-built diagrams with a high degree of accuracy and precision using the global positioning system (GPS) and other surveying technology. Planning personnel use GIS to integrate geospatial data from all parts of LANL for use in the siting of new facilities. Operations personnel have the roles of data providers, data managers, GIS users, and sometimes customers.

Much of the operations data are generated and used in CAD systems along with associated tabular data. Typically CAD data are migrated to GIS format for producing maps and performing analyses. Infrastructure data are foundational (base or core information), of value to all users that focus work on LANL land (researchers and emergency responders alike), and therefore data documentation, currency, and accessibility are critical.

Environmental Monitoring and Restoration. The environmental monitoring and restoration personnel are responsible for the ongoing monitoring and environmental cleanup efforts at LANL. They generate massive databases of environmental conditions that are typically updated at frequent intervals and tied to modeling of contaminant transport and environmental effects. A diverse set of environmental experts focus on issues concerning waste sites, terrain, flora, fauna, soil and underlying geology, surface water, groundwater, and the potential or actual release of contaminants. They provide technical expertise and guidance for operational and planning purposes at LANL in addition to producing reports for regulatory agencies, government officials, and the public.

The varied sampling and analysis projects undertaken by environmental professionals require detailed topographic information and complex feature-rich geospatial data stored in large databases. Environmental GIS experts generate new geospatial data as well as depend on institutional databases (potential release sites, facilities, utilities, etc), and therefore they play all roles in the enterprise GIS (data providers, data managers, GIS users, and customers). They need advanced GIS software and high-end hardware, including off-site backup. As a result of the considerable time and money invested in local GIS capabilities, these workers may be concerned about the loss of data ownership in the process of evolving toward an enterprise GIS.

Research and Programmatic Work. Scientists and engineers performing programmatic work and research are involved in a broad range of activities, often with comparatively specialized scope and focus. Researchers typically use a suite of basic geospatial and tabular data (topography, climate, geology, ecology, etc), and derive specialized datasets and results from numerical models, and they may play all roles in the enterprise GIS. Typical research projects are carried out by a small number of people, and the scope of the work may focus on the LANL region, or on external sites. The product of research is often a report or professional publication for a national or international audience, rather than exclusively for internal LANL use. As a result, the need or interest for data sharing or exchange may be less for researchers. However, as a group, researchers are heterogeneous in their activities and perspectives. Research focusing on the LANL area may rely heavily on institutional data resources, from topography to soils to infrastructure, and the product of the work may be of interest to many others within LANL. Whereas, research focused on other regions may be of less immediate institutional concern.

Historically, most research and programmatic datasets have a relatively short lifetime, are generally stored locally, and are eventually superceded or abandoned once the analysis results are published. As a result, there has been small need for off-site backup, although research projects might utilize it if it were available, especially for research that focuses on the LANL region. Research geospatial data that may be of value to others at LANL are often not documented or captured. Increasingly, researchers are working with large, georeferenced datasets, and funding agencies, including DOE, are requiring more attention to integrated information management, including requiring researchers to document and make data sets available to others. As multidisciplinary and long-term research increases, and as researchers require more data fusion and long-term study, there are increased needs to document, archive, and make data accessible. These trends are expected to continue.

Many LANL researchers may not have an interest in enterprise GIS because of the perceived large burden of effort and money on their short-term projects to make data compatible (format, metadata, QA) with a centralized system for which they may see little immediate use. Enterprise GIS may be viewed as either a luxury that researchers cannot afford or as a burden that takes resources away from research efforts. This perception is especially strong in the case of research that focuses on geographic regions external to LANL. There are also concerns about data ownership and sharing of incomplete or unpublished ideas with competing research groups. On the other hand, enterprise GIS can permit researchers the opportunity to "publish" the results of their work in a way that is not possible with traditional paper media. Large geospatial datasets can be stored in a geospatial data warehouse and "published" via an Internet map server to allow others to explore the richness of the research results that are referenced in a traditional research paper.

Because a significant set of LANL research and programmatic work focuses on solving institutional problems (e.g., models of contaminant transport, design of analytical technologies, and analysis of land stewardship options), GIS needs can often carry an institutional perspective that strongly overlaps the interests of managers, operations, and environmental monitoring/restoration stakeholders.

Rationale for Enterprise GIS

This analysis of major CGRP-GIS stakeholders reveals several fundamental problems for implementing an integrated GIS to support fire rehabilitation efforts institution-wide. First, the stakeholders are working with different deadlines and different goals. Environmental monitoring and infrastructure reference databases (e.g., utilities, structures, roads) emphasize longer-term institutional and regulatory concerns, such as change control, updates, consistency of data format, and documentation of data sources. Research projects often place a higher priority on short-term goals of rapid analysis and publication, with much of the knowledge held in the minds of individuals. Data documentation, consistency of data format, and long-term archiving may not be a high priority. Once a given research project is completed and summarized, data are stored in an ad hoc fashion and often are eventually lost.

However, these disparate approaches have many common needs, such as data quality standards and a geospatial information management plan. With proper design, enterprise GIS can promote data sharing while protecting data security and while promoting increased integration of operations and research efforts for the benefit of the institution. Enterprise GIS can enable all stakeholders to work more for the good of the institution, by helping project and operations workers to provide complete and accessible data, by helping researchers to conveniently complete neglected tasks like documentation and archiving, and by ensuring that research results are available to serve LANL management, operations, and environmental monitoring/restoration. Large and small projects alike can benefit from metadata, consistent data formats, and data archive and backup arrangements. Long-term operations work, for example, can benefit from the results of specialized studies of slope stability or wildlife habitat, while readily available infrastructure data may be of great value to researchers. These connections must be highlighted in the context of enterprise GIS with efficient procedures that do not add undue burden on already overtaxed project teams. Otherwise, "different user and task requirements will continue to result in different systems being established. Different systems result in incompatible data formats and restrict the flow of information and data exchange between users" (Information Architecture Project 1997:5).

The actual form of enterprise GIS at a large institution like LANL may fall anywhere along a continuum of centralization. At one end of the spectrum, data and metadata are stored in a centralized repository available to all users. At the other end, individual GIS users maintain local datasets, and the "enterprise" aspect is limited to standardized software and policies on metadata, data quality, standardized formats, and data sharing. Several models for enterprise GIS at LANL can be developed from this continuum, based on the diverse needs of the GIS community:

1. Centralized data repository. All geospatial data (and pertinent tabular data) are entered into a LANL-wide data repository. Data stewardship and responsibility for data currency and accuracy remains with the data owner/generator. The repository is administered by a dedicated team. A central metadata clearinghouse also resides with this group. Enterprise-wide data standards ensure data quality, format, and documentation (metadata). This format provides assurance of uniformity in standards of quality, documentation, and format, but it requires significant administrative overhead for maintenance of the repository, and it requires that individual data generators invest the time to follow all data quality and format procedures and coordinate closely with the repository administrators on issues of data currency and change. A major advantage is that centralization can ensure adherence to policies and standards, and ensure data availability, security, and completeness.

2. Distributed data storage with enterprise standards. This middle-ground option calls for only as much centralized storage of geospatial data as is necessary for programmatic or operations needs. For example, a centralized data warehouse could supply the EOC with a complete warehouse of important facilities, utilities, topographic, and environmental data crucial to emergency management; other data not required by the EOC need not be included in the central repository. Many datasets could remain in the stewardship (and storage) of the data owner, but enterprise standards for data quality, metadata, format, and access could allow efficient data sharing. The only uniformly centralized aspect of this model is a metadata clearinghouse. From the information in this repository, a user could contact the data owner for more details and access information. In this model, the responsibility lies clearly with the data steward to properly document, archive, and provide access to data.

3. *Ad hoc* data sharing. This model is essentially *status quo*, with individual working relationships dictating the ease and efficiency of data sharing. Users develop informal understandings among themselves about the existence, accuracy, currency and format of individual data sets. This format places the minimum burden on individual GIS stakeholders and keeps the responsibility of data stewardship with the data owner. However, overall data sharing is inefficient, and the possibility of redundancy and duplicated efforts is great. This model allows for possible enterprise policies for metadata, format, and quality, but it is up to the individual users to follow the standards and it is the least efficient and effective on an enterprise scale.

All GIS stakeholders at LANL share a common challenge as each works to assemble the necessary geospatial and tabular data for a project at hand. Data gaps present the quandary of spending money to reproduce data that probably exists somewhere at LANL versus spending the time to hunt it down, along with the necessary quality control and source information (metadata). Once the data are located, further hurdles include data compatibility, currency, accuracy, and access. Is the data owner/steward someone with whom a working relationship exists? How is this communication gap bridged? The next section delves more deeply into these and other issues and offers potential solutions in the form of a prototype enterprise GIS.

ENTERPRISE GIS

The Ideal: An Integrated View of Geospatial Information Management

A complete, or unbroken, cycle of geospatial IM involves flows of data from source to database, from database to data user, and, if modifications have been made, from the data user back to the database, with necessary steps to ensure that data is complete, secure, documented, and accessible (Figure 3). These steps include a suite of necessary data operations: formatting, quality control, documentation (metadata), cataloguing, tracking, backup, delivery, and updating.

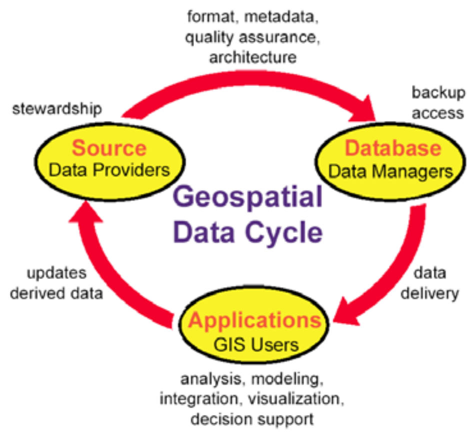

Figure 3. Cycle of geospatial information management. "Source" refers to original or modified data as provided by the data steward.

The realm of geospatial data has three main aspects: source, database, and application. At the source of geospatial data, stewardship is the generator's primary responsibility, including ensuring quality and accuracy; performing updates and corrections as needed; and protecting the resources from accidental loss or security violations (Information Architecture Project 2001). The data steward is the primary point of contact for anyone accessing the resources. Data documentation (metadata) must be part of the data steward's effort during data generation. Data generation has generally been the purview of various facilities and operations organizations at LANL, and it is always part of the science (e.g., research and project support) process. However, stewardship has often been minimized by budgets and not given the importance it deserves, except in some specific areas of data collection, e.g., remote sensing and geologic mapping.

Data management involves data organization, access, delivery, and change control (the tracking and management of updates and edits to geospatial data). This area typically has been the realm of information sciences professionals—R&D information science, professional data managers and database administrators.

Most of the effort in GIS to date has been in the "application" area because of the need for immediate solutions to customers' problems. Often this area involves special-case applications tailored for individual projects, rather than generic applications of more general utility. Traditionally applications developers have not considered the relationship of end solutions to information sources and management. While these aspects blend together in varied measures on any given project, they are typically treated separately, without the benefit of an integrated geospatial data management plan. As geospatial information management matures at a large institution like LANL, there is an integration of all three aspects.

Data sharing is key for smooth and efficient operation within the enterprise GIS environment, enhancing efficiency, effectiveness, and decision-making ability (Pinto and Onsrud 1995). Given the high cost of geospatial data acquisition and maintenance, effective cataloging and sharing of the data enhance the collaborative development of geospatial data resources and can reduce overall costs. Since the earliest studies of enterprise GIS at LANL, there has been an understanding of the basic elements of effective data sharing:

> "…a central repository where users can go to find the most appropriate data; metadata that provides the information to evaluate the suitability of the data to the task at hand; stewardship of the data that delegates responsibility of maintaining the data and its metadata; and data transfer standards that provide for the expedient and efficient dissemination of that data" (Information Architecture Project 1997:4).

The importance of quality control and proper metadata for geospatial data cannot be overstated. Quality control involves standardized procedures for ensuring that data meet standards for geospatial accuracy and reliability of associated attributes (data fields). "Not only must agencies sharing data trust each other's quality control procedures, but they must also be able to communicate knowledge of data quality—sources of error, types of imperfection, measures of accuracy—in an objective and useful way" (Goodchild 1995:414). This "useful way" is metadata, or data about data. Metadata serve to document essential characteristics of data, including source, content, geospatial extent, format, quality, means for access, etc. The Federal Geographic Data Committee (FGDC), under presidential order (Clinton 1994) and with representation by DOE, has formulated national standards for metadata and requires that all federal government agencies produce metadata for their geospatial data (FGDC 1998, OMB 2002). Compliant geospatial metadata enable construction of searchable catalogs of geospatial data and serve as the "institutional memory" to ensure that data are available in the future.

The necessary policies for data quality and documentation must be standardized from the individual project teams to the level of the institution. Without standards for data quality and metadata, there is little hope for efficient data sharing at an enterprise level. Heterogeneous (or nonexistent) geospatial information management policies result in inconsistencies in data quality and metadata, while "data that meet the quality standards and needs of one agency will frequently fail to meet the different, or more exacting, needs of others" (Goodchild 1995:413). An institutional geospatial information management plan integrates GIS efforts by specifying policies and standards, procedures to implement these policies, the infrastructure needed to meet geospatial data needs, and which data are key and how they can be accessed.

Challenges for Enterprise GIS at LANL

At LANL, as in many large institutions, the level of expertise of individual staff and the capabilities of project teams is high. Because of the large size and long time frame of various projects, there is also significant GIS infrastructure in place in the form of hardware, commercial software licenses, and custom software applications. There is typically good informal communication among GIS experts in the institution as a result of ongoing collaborations. In general, small teams of GIS professionals throughout LANL work well internally and meet project needs within the scope of individual projects and organizational mandates.

However, at the larger level of the institution as a whole, the cycle of information management is broken, and GIS coordination is difficult. The translation from small, semi-independent GIS teams to institutional, enterprise GIS is faced with many challenges, including duplication of facilities, lack of coordination, incompatible data format and architecture, inconsistent quality control and change control, and lack of data protection. Understanding these challenges is key.

In contrast to the idealized cycle of geospatial information management (Figure 3), the constellation of semi-independent GIS teams within LANL constitutes an inefficient institutional system, without consistent geospatial IM policies and procedures. Because GIS capabilities have developed independently in many parts of LANL, the institution suffers the hindrance of "stove piping," in which there is a lack of coordination and duplication among these facilities (Fletcher 1999) (Figure 4). While many GIS experts at LANL enjoy collegial relations and share information and experiences, the disadvantage of this informal organization is that access to geospatial data depends on personal relationships and not necessarily on policy. GIS efforts developed for individual projects result in many different standards, levels of quality, and means for data access. In addition, these independent GIS facilities have developed unnecessarily redundant infrastructure for data management and delivery, and insufficient effort is focused to define the distinct institutional role of each GIS facility.

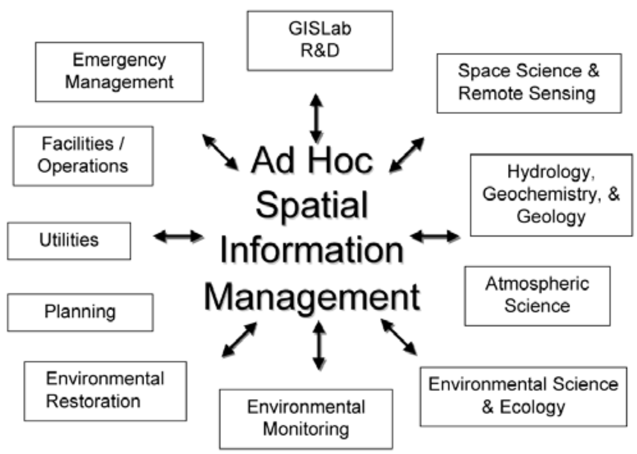

Figure 4. Organizational divides produce a lack of coordination among LANL GIS facilities

The availability of high-quality geospatial data depends upon sufficient funding, staffing, and recognition of the institutional value of efficient access to such data. In many cases, lack of funding and insufficient staffing within teams that manage key geospatial data (infrastructure, environment, operations) results in poorly documented and poorly organized information that may not be available to emergency managers during a crisis. Inaccuracies in key geospatial data are common, but they often go uncorrected due to the lack of available staff. When geospatial data do meet quality standards, staff do not have time (or sufficient obligation) to prepare basic metadata, often because the data are collected and used primarily for short-term projects. Data providers may forget important details, and key personnel may leave the institution. The result is inability to locate data, uncertainty about limitations and quality, and the possibility for incorrect use of data.

The lack of a central metadata catalog for and access to key institutional geospatial data can produce the failure of a GIS during times of greatest need. Complete and reliable geospatial data were not readily available during the May 2000 Cerro Grande wildfire or in the post-fire rehabilitation effort. Data users frequently waste time and resources trying to locate geospatial data because data reside at many locations and users have no efficient means to find sources. While this problem is particularly acute during crises, more routine program goals cannot always be met due to unavailable, inaccurate, or conflicting data. There is no guarantee of current and authoritative data, fully compliant with standards. The result of this missing documentation is a lack of a comprehensive institutional inventory of available geospatial data and their origin and quality. Without this inventory, a master list of geospatial data needs, based on a comprehensive analysis of data gaps and inadequacies, cannot be developed.

Given these critical needs, a central LANL geospatial data repository can ensure that geospatial data are in standard formats, that quality control has been performed, and that data are available without extensive searching. Such a repository can only be effective if it is supported at an institutional level and if policy requires that key geospatial data be placed in the repository. Establishment of an institutional geospatial data warehouse brings institutional benefits for day-to-day operations, programmatic needs, research, regulatory functions, decision-making, and emergency management.

A final challenge for institutional GIS is the elimination of duplicate or redundant GIS infrastructure and the completion of the enterprise system. Repetitive and redundant software license agreements between each project team and a central software provider can result in much higher costs than is possible in a coordinated (institutional) licensing agreement. While each team requires its own hardware, an institutional perspective can provide the benefit of coordinated, off-site data backups. The additional challenges of geospatial database management and access via the web (internal and external) can also be minimized through coordinated efforts at the institutional level. In conclusion, a well-formulated enterprise GIS design is good business.

Nowhere is the need for enterprise GIS better illustrated than in the case of the EOC at LANL. As the difficult experience with geospatial data management during the Cerro Grande fire demonstrates, emergency managers require current, accurate, and accessible geospatial data for decision support during times of crisis. The sources for these necessary data encompass nearly every part of the institution, necessitating the active participation of data stewards from the operations, project, and research communities.

Although the members of the EOC represent many areas of LANL operations (security, material safety, etc.), the organization and structure of the EOC at the time of the May 2000 Cerro Grande wildfire reportedly was not well-suited to dealing with natural hazards (Salazar-Langley et al. 2000a). As a facility primarily designed to deal with radiological and security incidents, the magnitude of the wildfire and its far-ranging effects were outside the original scope of the EOC, although the emergency managers during the fire were able to rise to the challenge. Emergency managers reported that their response during the wildfire was limited by unreliable electrical power and networked communication and a lack of accessible, up-to-date geospatial data during and immediately after the fire (Salazar-Langley et al. 2000b, Keating et al. 2001). The emergency managers' decisions were limited to information (spatial and tabular) on-hand in the EOC, even though it was often not the most current or accurate data available at LANL under everyday (networked) conditions. Real-time data on meteorological and fire conditions and status of facilities and personnel were difficult to access from the EOC. Incompatible devices (radio, wireless, landlines) made communication difficult among field workers from multiple agencies who needed to share data. The lack of an efficient means for publicizing status messages to the evacuees and interested public resulted in a perceived information vacuum (Salazar-Langley et al. 2000a).

Model of an Enterprise GIS: GISLab

Each of the challenges and institutional deficiencies described in the previous section must be recognized and addressed before an effective geospatial information management plan can be formulated and enterprise GIS can be properly evaluated and perhaps implemented at an institution such as LANL. As a first step toward designing an enterprise solution to meet the needs of the institution, a model (prototype) enterprise GIS was developed to serve the CGRP. The GISLab team at LANL was tasked with designing and implementing a GIS in response to the needs of the CGRP effort by gathering geospatial data related to the Cerro Grande fire and the subsequent recovery and rehabilitation efforts. These data are stored in a geospatial data warehouse for efficient access by stakeholders via network connections or a simple web interface. Additional new technology will provide programmatic and operations staff more rapid access to data and increased functional and analytical capability. This enterprise GIS will assist scientists and emergency managers in anticipating long-term forest fire effects and in dealing with future emergencies in the Los Alamos area. This effort illustrates key concepts of one enterprise GIS design that have been effective solutions, as well as the solutions that remain elusive. A more complete description of the GISLab information architecture and technical GIS solutions is provided by Witkowski and others (2002, 2003).

Data Management. At the center of the CGRP-GIS is a geospatial data warehouse with associated data policies and procedures. The flow of geospatial data in and out of the GISLab geospatial data warehouse is cyclic (Figure 3). At the source, the data provider produces a version of the data in a standard format, determines its quality (accuracy and precision), and documents it in the standard metadata fields. This geospatial data is then added to the architecture of the database, and it is managed according to standard policies and procedures (including backup, access control, etc.) At any time, this data can be used in visualization, analysis, or modeling, which may involve modification through updates, edits, or additions or the generation of derived data (e.g., floodplain maps from DEM). The cycle begins again as the modified geospatial data are entered into the database with revised attributes and metadata. For the data warehouse to function effectively this cycle must be preserved through the understanding and participation of all institutional GIS users. The GIS breaks when any part of the cycle is bypassed or when data leaves the system.

Proper management of geospatial data provides infrastructure for 1) customer support – receiving and tracking of data from data providers, 2) staging – preparation of data for placement in a data warehouse, 3) data warehousing – archiving and management of data, and 4) data delivery – via portable media, networks, or the Internet (Figure 5). The GISLab information architecture and data management practices ensure predictable customer interactions (both data providers and data users), quality control and documentation for data in the repository, standardized GIS interfaces and tools for access and analysis, and data security by accommodating change control and access restrictions as specified by the data providers.

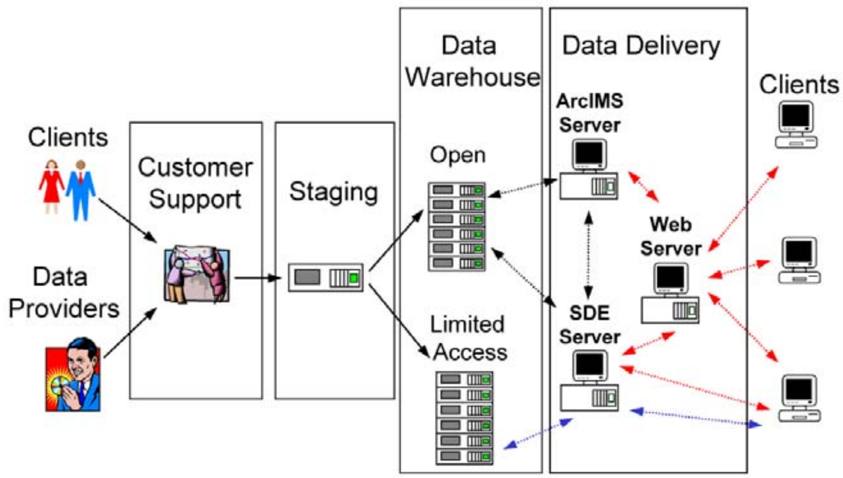

Figure 5. Data management in the GISLab Geospatial Data Warehouse.

Clients (data users and providers) contact the people and resources of GISLab through a "customer support" interface, where needs are evaluated and requests are entered using a web-based interface. In the staging area, new geospatial data are processed before placement in the data warehouse. Data entering staging are archived in their unmodified source format and the initial quality control check is performed. Standard procedures (Figure 6) are applied to assign names for data layers, to ensure complete and consistent metadata, and to provide for access control when required by data providers.

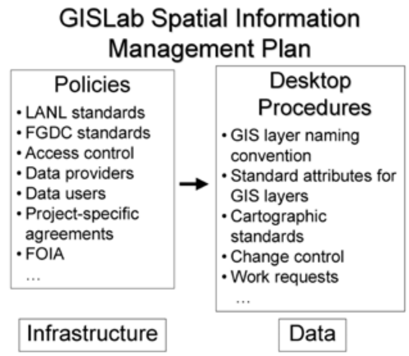

Figure 6. GISLab geospatial information management plan.

Metadata is checked for compliance with FGDC standards. Once the new geospatial data have been processed, they are transferred from staging to the data warehouse. The major elements of the physical data architecture are illustrated in Figure 7.

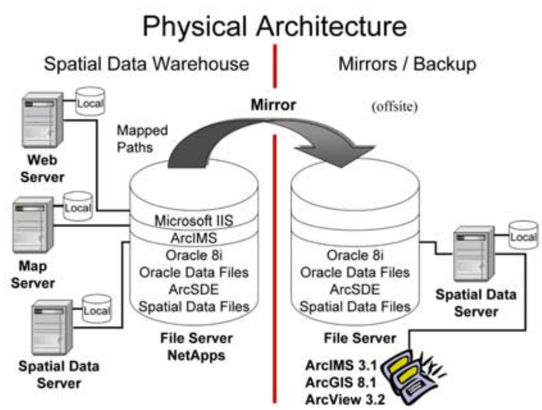

Figure 7. GISLab physical data architecture, including offsite backup

The data in the geospatial data warehouse are accessible through three separate means. First, internal LANL users may access files via a GIS software client (e.g., ArcView, ArcGIS). Second, GISLab (team) users may access files directly using file sharing across the UNIX and NT networks. Finally, users without GIS client software may view selected data layers or maps via the Internet using a web browser. Internet Map Service (IMS) capability provides the means to have different content for internal (LANL) and external users.

Tools and Services. In addition to the core hardware, database, and policies and procedures for geospatial data access and use, GISLab staff have developed tools and applications to streamline the information management process. These include a web-based interface for browsing the metadata catalog, entering GIS service requests, consulting policies and procedures, and interacting with members of the LANL GIS community.

The GISLab website (http://www.gislab.lanl.gov) provides a platform for institution-wide access to the CGRP-GIS. The site contains pages that describe the geospatial data warehouse, data quality and metadata policies and procedures, and data access agreements. An online catalog of metadata describes data sets residing both in the GISLab geospatial data warehouse and elsewhere at LANL, analogous to a central "card catalog" for geospatial data. Maps are made available for direct download (LANL intranet) via a virtual Map Gallery. The website serves as a portal to several GISLab tools, including the request system for entry and tracking of requests for GIS services (cartography, applications development, geospatial analysis, etc.). The website is hosted by an external web server that makes public much of the general information about enterprise GIS work at LANL, while material suitable for internal use only is served by a linked server behind the firewall. The GIS community at LANL is kept informed by pages containing information about stakeholders (data users and providers) at LANL and a link to the LANL GIS User Group home page.

The CGRP-GIS website (http://www.cgrp-gis.lanl.gov), linked to the GISLab website, provides a clearinghouse for CGRP-related geospatial data. The data catalog section of this site provides direct downloads of the Burned Area Emergency Rehabilitation (BAER) Team geospatial data layers and maps in shapefile and ESRI export file formats. An IMS capability on this site allows a user with a web browser to view sets of data layers that have been compiled as maps, and it provides limited GIS capabilities. A feedback page provides a link to the consensus project for the CGRP-GIS, which seeks to identify and clarify conflicts and reach consensus among stakeholders regarding issues of GIS design (Keating et al. 2001). Like the GISLab website, the CGRP-GIS site is hosted on an external web server, with links to selected pages that reside behind the firewall. In this way, we provide public access, while providing the capability to restrict access to LANL-only information.

Access to geospatial data stored in the GISLab Oracle database is enhanced by the use of Spatial Database Engine technology (ArcSDE, ESRI, Inc.). The software is installed on the database servers and plays a fundamental role in a multi-user, enterprise GIS by providing the business rules for accessing and storing geospatial data within a database management system. For GISLab data managers, it provides a framework to organize, manage, and store geospatial data. For GISLab customers using various client software (e.g., ArcInfo, ArcEditor, and ArcView) it allows simultaneous access to our geospatial data from different physical locations. For large data sets, SDE enhances client performance by serving a subset of the data depending on the user's area of interest, thereby eliminating the need to download massive data files to the client workstation. ArcSDE is also used to serve geospatial data across the Internet via ArcIMS map services.

The Internet Map Server (IMS) capability is in use for the CGRP-GIS and other GISLab clients, and its utility is expected to increase over time. This technology allows users to view and perform limited GIS functions on geospatial data layers via an Internet browser interface, without the need for GIS software installed on the client computer. Map Services are compiled from related geospatial data layers (basemap, image, and feature data) and published as an image-based web site for intra- or Internet access. IMS provides an efficient means for a wide range of stakeholders to view geospatial and tabular data, as well as metadata, related to the CGRP effort. Map Services can be built and modified rapidly to meet project needs, providing a means for near real-time information dispersal to stakeholders and decision makers. Web technology allows rich density in the data available on a single map via database queries (and associated tabular reports) and embedded, hyperlinked images and documents. This technology has the potential to greatly enhance decision support for natural hazard mitigation and emergency management.

The GISLab Request Tracking System is a web-based application that enables team members to manage numerous monthly requests for GISLab services. The requests are made both internally by LANL groups and externally by a variety of agencies, and the Java-based application stores request information in the GISLab Oracle database. Each request is tracked from the time the request is entered to the time the request is completed, and it is retained in the database to provide historical information. Query and reporting interfaces allow users and managers to assess and document the status of individual requests and groups of requests, and to gather statistics on the requests. In the case of ArcGIS map-production services, the request system is capable of importing map properties directly produced by GISLab map production tools. The integration between map production and request tracking streamlines the flow of information through the GIS enterprise.

Metadata are applied to geospatial data layers using ArcCatalog (ESRI, Inc.) and custom extensions to this product, including MetaBatch developed by GISLab. The GISLab MetaBatch enables us to rapidly apply metadata to large sets of data layers. In addition, the GISLab can track feature-level metadata through the use of standard data fields within each geospatial data layer. These commercial and custom GIS tools enable us to publish metadata documents on the GISLab web site, providing a resource that can be browsed and searched to query and retrieve information about geographic data layers. Hierarchical indexes (HTML), interactive ("clickable") map indexes, and IMS metadata server tools (ArcIMS) are in development for access to metadata. This resource forms a cornerstone for enterprise GIS at LANL.

The GIS community at LANL is supported by several tools and applications. A database contains information that pertains to GIS stakeholders, including data providers, GIS professionals, and a diversity of customers. Information retained in the database includes contact information, specific areas of interest and expertise, hardware and software platforms used, and information about GIS service requests. The stakeholder database is an important tool for establishing a cohesive network of data providers, GIS professionals, and end users, and it helps unify the GIS community and can facilitate collaboration. In addition, an internal LANL GIS user group and email listserver have been established to allow professional GIS users to interact. GIS-related issues can be posted and discussed in an open-forum environment that helps users to reach solutions with the aid of the GIS community, as well as keeping members informed about new technologies and upcoming events.

Metrics of Success. The success of an enterprise GIS can be measured according to performance and service. Specific performance metrics include availability of data and other resources, reliability of the system, and system usage; service metrics include funding stability; productivity, and degree of data sharing and collaboration (Witkowski et al. 2003). The enterprise GIS designed by GISLab was required to ensure a complete data cycle via the design and implementation of a data management system with data access policies, consistent standards and procedures, efficient storage design, and multiple access methods. The GISLab model can be considered a success in its implementation of a sound data management design (data warehouse hardware, software, and policies and procedures). System reliability (e.g., adherence to metadata and data formats, procedures, standards, and policies) and availability (e.g., system downtime below 20 hours per year) were good. Usage showed consistent growth in terms of numbers of new users and connections to the data warehouse. Funding was stable for the life of the CGRP; continued institutional funding for the GISLab prototype is uncertain; and enterprise GIS is moving to a larger scale at LANL, as described in the next section.

Scaling Up: Institutional Progress

The enterprise GIS methods developed for GISLab and the CGRP have broader applicability at the level of the LANL complex as a whole (Witkowski et al. 2003). Progress is being made to scale up these ideas via institutional working groups and the development of a new EOC. In Fall 2001, the LANL Chief Information Officer (CIO) formed a GIS Working Group, charged to summarize the current status of GIS at LANL, identify key issues concerning institutional GIS, envision a viable future state for GIS at LANL, and offer recommendations for the implementation of enterprise GIS. In January 2002 the Working Group completed their work and reported to the CIO Policy Board. This group identified fifteen key issues concerning GIS for LANL, addressing data inventory, stewardship, and sharing; metadata; software licensing; funding; GIS roles; and the perceived window of opportunity for change in institutional GIS, among others. Based on the recommendations of this group, two GIS Steering Committees (Management and Technical) were formed in spring 2002 to develop a LANL geospatial information management plan based in sound business practices, explicit roles and responsibilities, and institutional geospatial information standards and policies. As cross-organizational groups, they will enhance lateral interaction and build connections, trust, and buy-in among the GIS community (Pinto and Onsrud 1995).

The design of new EOC, planned as part of the post-fire rehabilitation project at LANL, provides solutions to the limitations observed during the Cerro Grande wildfire. First and foremost, all necessary geospatial and tabular data will be available to emergency managers from file servers, potentially to be located in the EOC, ensuring data self-sufficiency in concert with the planned self-contained electrical power. These file servers will receive snapshots of critical institutional data on a regular basis to maintain data currency and accuracy. In addition, off-site backup facilities will maintain a copy of the EOC data repository—as was the case during the Cerro Grande Fire, access to external (e.g., out-of-state) data sources may be possible even when access to databases within the LANL system is not. This data repository concept requires that institutional data not only are accessible, but also that they are in standardized or compatible formats, including CAD, ArcGIS, Oracle, SQLServer, and other tabular and geospatial formats; this may pose the greatest design challenge. Field GIS systems on hardened laptops will provide integrated data resources for emergency management personnel in the EOC and in the field (Hart 2001).

DISCUSSION

Perspective on Evolution of Enterprise GIS

The use of enterprise GIS is a natural result of the evolution in geospatial data sharing within institutions, but this change can be painful. A natural part of evolution is resistance to change, and this is manifested in unique ways at each institution. This resistance is affected by different stakeholder roles and stereotypes (e.g., operations vs. research). In addition, the typically excellent working relations among GIS professionals at the grass-roots level can be limited by organizational divides. In the final analysis, an enterprise GIS design for an organization like LANL must meet the needs and missions of a broad spectrum of stakeholders; the challenge lies in striking a balance in the degree of centralized storage, administration, and procedural control while serving the needs of the GIS community for streamlined data documentation, access, and compatibility. Beyond the sharing of geospatial data, the enterprise facilitates sharing of information and GIS resources as well.

Role of CGRP

The Cerro Grande fire highlighted the need for institutional GIS solutions and presented a new opportunity for improved data sharing in an enterprise setting at LANL. The CGRP effort provides a common project with goals shared by the operations and research communities. Geospatial data is necessary for basic project planning and execution, and derivative data is being produced abundantly. A spatial data warehouse for the project is necessary, but obstacles to centralization of CGRP data (storage, access, quality, documentation, format, and architecture) were not overcome until long after most CGRP-related efforts were well underway (infrastructure rebuilding, flood mitigation, forest thinning, etc.). The challenges to a CGRP-wide GIS encountered during this work laid the foundations for expanding the scope to a LANL-wide GIS organization, and the institution as a whole will benefit from the CGRP-GIS efforts. However, as the scope for integrating GIS activities at LANL has expanded, so has the complexity of design and implementation to meet stakeholder needs.

Enterprise GIS and Natural Hazards Mitigation

Following the Cerro Grande fire and the immediate recovery activities, GIS has been used in a variety of hazard mitigation activities at LANL, including forest thinning, forest fuel characterization, floodplain delineation, flood and erosion modeling, and wildfire modeling. Many of these post-fire GIS activities have been limited by the lack of enterprise geospatial data management. From basic cartography to complex spatial analysis, these efforts require efficient access to current, accurate data layers, as illustrated by the case of planning for forest thinning for wildfire mitigation. GIS specialists produced maps of proposed areas of tree cutting on LANL land using analyses that incorporated buffers around electrical lines, cultural and archaeological sites, and core habitat for threatened species. Additional factors, like slope and vegetation class were also important for determining thinning prescriptions. This work was done rapidly on short deadlines and required access to accurate, current, data layers of known lineage. Well-documented data provides confidence in analysis and reporting, the latter being especially important for large-scale hazard mitigation work. The incomplete nature of the institution's metadata and data access infrastructure provided frustrating obstacles to these tasks and highlighted the need for an institutional geospatial data management plan.

Enterprise GIS can also provide enhanced hazard mitigation tools in the form of stored physics-based model results. Results of flood and erosion modeling can be used to plan sediment and flood water retention systems as well as to aid in evacuation plans during a crisis. Predictive model results (fire, flood, atmospheric dispersal, earthquake, volcanic eruption) could be useful to emergency managers before and during a crisis if the results are incorporated as layers into a GIS that supports the evaluation of possible emergency scenarios—a spatial decision support system. Static maps provide an absolute minimum of information, but a GIS with access to important institutional data and model results can provide a much more powerful hazard mitigation and emergency management system.

Resistance to Enterprise GIS

Substantial steps to overcoming inefficient geospatial data management are relatively easy to identify but can be difficult to implement. An outside observer recently noted that the “soft-money culture” at LANL manifests an environment in which strong competition for perceived limited resources leads to many isolated, independent, project-centric GIS groups with limited motivation for data sharing and collaboration (Glazer 2001). Institutional resources are poorly allocated in redundant hardware, software, and data generation, an even bigger problem when the idea of a central data repository is raised. This culture leads to a basic lack of will to cooperate when organizational or personal incentives to share are insufficient to overcome the impediments (Craig 1995). Data and information are held closely as sources of control and power and there is an unwillingness to share data that is viewed as “proprietary” (Pinto and Onsrud 1995). One result of uncoordinated teams in large institutions is “stove-piping,” or the redundant development of complete GIS capabilities and duplicate data, forming small, self-contained (and often competitive) pillars of activity.

These problems are often exacerbated by the apparent division between the needs and goals of facilities/operations and project/research staff. The differences in geospatial data management posed by long-term operations and monitoring work versus variable-length research projects may be difficult to overcome. Support for enterprise-level data stewardship (documentation, change control) may or may not be provided in operations budgets, and individual researchers do not typically budget for IM, especially for smaller, short-term projects. A divide commonly manifests in a difference in data formats between the Computer Aided Design (CAD)-based systems in operations and the spatial-analysis software (e.g., ArcView and ArcGIS) used by researchers.

The state of GIS at LANL is the product of years of institutional history, and the process of change must begin with characterization of the current state of the art in the institution. This process highlights problems and limitations, especially those identified during crises like the Cerro Grande fire. Despite the extraordinary efforts put forth by many individuals in the aftermath of the fire to verify the safety of facilities and reopen LANL, the fire recovery effort highlighted many problems in the institution’s IM system, including the lack of an institutional emergency plan and sometimes-poor coordination between divisions. Stove-piping of resources and efforts was highlighted as a problem resulting from this poor coordination and leadership. Problems in coordination of non-centralized efforts resulted in redundancy, duplication of efforts, “turf guarding,” and political wrangling over roles and responsibilities that reduced the effectiveness of the recovery effort. Data-sharing was not effective. The synthesis of this post-fire analysis included a recommendation to integrate facility and programmatic response by evaluating interdependence and relationships (Salazar-Langley et al. 2000a, Salazar-Langley et al. 2000b).

Potential Solutions

What are the solutions to the institutional challenges to enterprise GIS and the identified root causes of reluctance to share data? Several steps have been suggested (Pinto and Onsrud 1995), beginning with integrating mechanisms such as task forces and cross-organizational teams, such as the CIO GIS Working Group and the new LANL GIS Steering Committees. Such efforts allow for lateral contact between organizations and a level playing field for the design of the enterprise GIS. Cross-organizational teams seek to identify "superordinate goals" (Pinto and Onsrud 1995:51) that transcend those of individual organizational units, provide value to all, and are attainable only through cooperation and data sharing within an institutional framework. In contrast to *ad hoc* personal exchanges, it has been shown that formalized processes and procedures for data exchange produce a better flow of information (Pinto and Onsrud 1995), due to a better understanding of individual responsibilities and expectations. Furthermore, each GIS stakeholder must have incentives to participate, ideally in the form of resources for unmet needs. These solutions can be implemented through the organizational process of the enterprise GIS.

Attention must be paid to the intangible aspects of the data exchange as well, including the real or perceived accessibility of key individuals and the quality of exchange interactions. Parties must feel that obligations will be met and promises kept, and there must be no real or perceived tendency to distort or hold back information, especially to leverage power or attain personal goals at the expense of others (Pinto and Onsrud 1995). These aspects of a climate conducive to data sharing are more difficult to include in organization or procedures, but they are the positive outcome of an open, fair enterprise GIS design.

There is a clear need for efficient data sharing and enterprise-wide data standards, but complete centralization of geospatial (and much tabular) data may not be in the best interests of the diverse LANL GIS stakeholders. The metadata clearinghouse must be constructed and populated, per Executive Order 12906 (Clinton 1994), as updated in OMB Circular A-16 (OMB 2002), and standards for data quality, format, access and documentation must be enforced by a cross-organizational body such as the LANL GIS Technical Steering Committee. While many individual data sets can reside with the data owners, the existence, status, and access mechanism must be made known. Certain core data of near-universal utility, such as LANL infrastructure, topography, and orthophotography, should be placed in a central repository with adequate change control and data currency administration. Implementation of an institutional solution to enterprise GIS requires slightly greater burden on individual GIS users, but the value in efficient data sharing far outweighs the extra work, especially as stakeholders adopt sound information management and business practices.

It is notable that several large government agencies have designed and implemented successful designs for enterprise GIS. These include the USGS EROS Data Center, NOAA, and FEMA, all of which use FGDC metadata to index and access datasets. These and other governmental agencies, including Sandia National Laboratory and the Hanford and Savannah River Sites in DOE, are grappling with issues of data stewardship, archiving, and future access (Bleakly 2001, Maryak and Bergameyer 2001, Rush 2001). Collaboration and exchange of ideas among data managers from these agencies may streamline the difficult process of implementing enterprise GIS at LANL.

Elements for Success: The Geospatial Information Management Plan

Jack Dangermond, CEO of ESRI, Inc. (2002) expounds five elements for success of enterprise GIS: 1) attain management support, 2) develop a plan, 3) be customer focused, 4) ensure in-house "ownership" for the process, and 5) build a "team of two" of technical expertise and management support to make enterprise GIS a reality. Several of these elements have been addressed by the cross-organizational CIO GIS Working Group, and work continues in the LANL GIS Steering Committees, especially the marriage of technical expertise and LANL management support. However, the key to success for enterprise GIS at LANL is the development of a sound geospatial information management plan.

Like any business plan, the geospatial information management plan must be financially viable and technically sound. According to Dangermond (2002), the plan should address five aspects: 1) definition and design specifications for enterprise GIS; 2) description of internal and external databases being managed; 3) plan for conceptual applications and database architecture; 4) system architecture, including hardware, software, and applications; and 5) implementation plan. The scope of the implementation plan encompasses tasks, methods, and activities; a schedule; funding sources; and organizational responsibilities. While the details are outside the scope of this paper, this outline provides the basis for the development of the geospatial information management plan.

CONCLUSIONS

GIS Stakeholders

The major GIS stakeholders at LANL, both for the Cerro Grande Rehabilitation Project and in general, include members of the facilities/operations and project/research communities. These workers represent data providers, data users, and customers of GIS products. Data needs, budgets, and timelines often differ. However, all of these stakeholders have common needs for effective institution-wide data exchange, including metadata cataloging and data accessibility, accuracy, currency, and compatibility.

Challenges and Solutions

The transition from numerous small, semi-independent GIS teams to an integrated, institutional GIS poses many challenges, including duplication of facilities, lack of coordination, incompatible data formats, inconsistent quality control and change control, and need for data protection. Potential solutions to the challenges to enterprise GIS at LANL are being developed by GISLab and others, including a CGRP spatial data warehouse with associated metadata clearinghouse, data policies and procedures, and standardized GIS interfaces and tools. The CGRP-GIS model could be used to serve the enterprise GIS needs of the larger LANL GIS community. Recent progress has been made toward institutional GIS solutions through the

LANL CIO. A LANL-wide GIS Working Group surveyed GIS capability, identified key issues, and offered recommendations. Based on these recommendations, two LANL GIS Steering Committees (Management and Technical) were recently formed to develop an institutional geospatial data management plan and to develop and implement geospatial data standards for the institution.

GIS and Emergency Management

Emergency managers require current, accurate, and accessible geospatial data for decision support during times of crisis. The sources of these necessary data encompass nearly every part of LANL, necessitating the active participation of data stewards from the operations, facility management, utilities, project, and research communities. In order to overcome limitations in infrastructure, IM, and communications, the plans for the new LANL EOC incorporate improved data management and communications facilities, including an onsite data repository (file server, database); integrated LANL database access and updates; offsite data backup; data visualization tools; improved communication tools; and mobile GIS. The decision-support system (DSS) must be fully developed to include capabilities for data integration and display (GIS), database access, model (scenario) evaluation, team communication, and information dissemination.

Enterprise GIS at LANL

Resistance to the development of institution-wide GIS is the result of several factors, including competition for limited resources and an apparent division between facilities/operations and project/research staff. These factors contribute to poorly allocated institutional resources, resulting in redundant hardware, software, and data generation. This occupational culture leads to a basic lack of will to cooperate and share “proprietary” data unless personal and organizational incentives are provided. Stakeholders must feel that obligations will be met and promises kept, and there must be no real or perceived tendency to distort or hold back information, especially to leverage power or attain personal goals at the expense of other organizations (Pinto and Onsrud 1995). The challenge for meeting the GIS needs of diverse teams and users in a large institution lies in striking a balance in the degree of centralization of data storage, administration, and procedural control while serving the needs of the community for streamlined data documentation (metadata), access, and compatibility.

The actual form of enterprise GIS at a large institution like LANL may fall anywhere along a continuum from a centralized data and metadata repository to distributed data storage with institutional standards concerning metadata, data quality, data format, and policies governing adherence to these standards and data sharing. A middle-ground solution that involves distributed data storage with enterprise standards appears to be the most likely, calling for only as much centralized data storage as is necessary for programmatic and operational needs. Under this institutional structure, data stewardship remains with the data owner, but enterprise-wide standards for data quality, metadata, and format allow efficient data sharing. The only uniformly centralized aspect is a metadata clearinghouse. Certain data of near-universal utility, such as infrastructure and topographic layers, should be placed in a central repository with adequate change control. The key to success for enterprise GIS at LANL is the development of a sound geospatial information management plan. Like any business plan, the geospatial information management plan must be financially viable and technically sound.

ACKNOWLEDGEMENTS

This work was funded by the Cerro Grande Rehabilitation Project – GIS effort and the Earth and Environmental Sciences Division at Los Alamos National Laboratory. We appreciate helpful discussions on institutional information management with Kathy Padgett, Gordon Besson, and Julie Crook. The manuscript benefited from thoughtful reviews by Fairley Barnes, Bill Carey, Michael Goodchild, Kendra Henning, Dave Howard, Randy Mynard, Joan Stockum, and three anonymous reviewers.

ABOUT THE AUTHORS

Gordon Keating is a postdoctoral research associate with GISLab, part of the Environmental Geology and Risk Analysis Group (EES-9) at Los Alamos National Laboratory (LANL). Gordon received his PhD in Geology from the University of New Mexico, an MS in Geology from Michigan Technological University (1991), and a BA in Geology from Carleton College (1988). He is building on a background in physical volcanology, hydrogeology, and numerical modeling of heat and fluid flow to address aspects of GIS and spatial decision support for mitigating natural hazards.

Paul Rich is the GIS Team Leader and a research scientist for GISLab at LANL. Paul received his MA and PhD in Biology from Harvard University, and his BA in Biology and Environmental Studies from the University of California at Santa Cruz. He has more than 15 years experience in GIS, remote sensing, environmental modeling, ecology, biogeography, software development, and information management design. Formerly, Paul was a tenured professor at the University of Kansas (1991-2001). Paul's interests include microclimate models applied to habitat characterization and biotic effects of climate change. Email: pmr@lanl.gov.

Marc Witkowski is the technical lead for GISLab at LANL and has over 9 years of experience in GIS, GIS code development, environmental modeling, scientific visualization, and enterprise GIS design. Marc received his B.S. in Earth Science from Eastern Connecticut State University. His current research interests are focused on the integration of numerically based predictive models and GIS for spatial decision support systems. Email: witk@lanl.gov.

REFERENCES CITED

Bleakly, D., 2001, Issues for Consideration for Long-Term Spatial Data Archiving. *Proceedings of 13th DOE Technology Information Exchange (TIE) Meeting*, Albuquerque, New Mexico, November 13, 2001.

Burrough, P.A. and R.A. McDonnell, 1998, *Principles of Geographic Information Systems*, Spatial Information Systems and Geostatistics (Oxford: Oxford University Press).

Clinton, W.J., 1994, Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure, Executive Order 12906. *Federal Register*, 59(71), 17671-17674.

Craig, W.J., 1995, Why We Can't Share data: Institutional Inertia. In Onsrud, H.J. and G. Rushton (Eds.), *Sharing Geographic Information,* (New Brunswick, New Jersey: Center for Urban Policy Research, Rutgers), 107-118.

Dangermond, J., 2002, Elements for Success: The Geospatial Information Management Plan, Personal Communication to P. Rich, February 5, 2002.

Davis, S.R., 1999, Enterprise-Wide GIS in Ohio. In von Meyer, N.R. and R.S. Oppman (Eds.), *Enterprise GIS,* (Park Ridge, IL: URISA), 31-43.

Federal Geographic Data Committee, 1998, Content Standard for Digital Geospatial Metadata, FGDC-STD-001-1998, http://www.fgdc.gov/.

Fletcher, D.R., 1999, The Interoperable Enterprise. In von Meyer, N.R. and R.S. Oppman (Eds.), *Enterprise GIS*, (Park Ridge, IL: URISA), 7-16.

Glazer, A., 2001, University of California Natural Reserve Board, Personal Communication to P. Rich.

Goodchild, M.F., 1995, Sharing Imperfect Data. In Onsrud, H.J. and G. Rushton (Eds.), *Sharing Geographic Information,* (New Brunswick, New Jersey: Center for Urban Policy Research, Rutgers), 413-425.

Gould, S.J. and N. Eldredge, 1977, Punctuated Equilibria: Tempo and Mode of Evolution Reconsidered. *Paleobiology*, 3(2), 115.

Hart, O., 2001, Update on Progress of Planning for New LANL EOC Information Management, Presentation to LANL GIS User Group, October 25, 2001.

Information Architecture Project, 1997, *Geographical Information Systems,* Los Alamos National Laboratory Information Architecture Standard IA-5A10-R1, LA-UR-96-471.

Information Architecture Project, 2001, Glossary of IA Terms, Los Alamos National Laboratory, May 2001, http://www.lanl.gov/projects/ia/about/glossary.shtml.

Keating, G., S. Rasmussen and M. Raven, 2001, *Consensus Building Tools for GIS Design*, Los Alamos National Laboratory Report LA-13894-MS.

Maryak, M. and J. Bergameyer, 2001, *Geospatial Database Software Requirements Specification,* Westinghouse Savannah River Company, Savannah River Site Report PEC-GIS-2001-0010.

Meredith, P.H., 1995, Distributed GIS: If Its Time is Now, Why is it Resisted? In Onsrud, H.J. and G. Rushton (Eds.), *Sharing Geographic Information,* (New Brunswick, New Jersey: Center for Urban Policy Research, Rutgers), 7-21.

Mynard, C.R., G.N. Keating, and P.M. Rich, 2002. *GIS Emergency Support for the May 2000 Cerro Grande Wildfire, Los Alamos, New Mexico, USA,* Los Alamos National Laboratory Report LA-UR-02-6746.

Office of Management and Budget, 2002. Coordination of Geographic Information and Related Spatial Data Activities, *Circular Number A-16,* Revised August 19, 2002, http://www.whitehouse.gov/omb/circulars/a016/a016_rev.html#2b2.

Oppman, R.S., 1999, An Organizational Model for Successful Enterprise GIS. In von Meyer, N.R. and R.S. Oppman (Eds.), *Enterprise GIS,* (Park Ridge, IL: URISA), 44-52.

Pinto, J.K. and H.J. Onsrud, 1995, Sharing Geographic Information Across Organizational Boundaries: A Research Framework, In Onsrud, H.J. and G. Rushton (Eds.), *Sharing Geographic Information*, (New Brunswick, New Jersey: Center for Urban Policy Research, Rutgers), 44-64.

Rush, S.F., 2001, Improving Access to Hanford Geospatial, *Proceedings of 13th DOE Technical Information Exchange (TIE) Meeting*, Albuquerque, New Mexico, November 13, 2001.

Salazar-Langley, C.A., D.L. Hall and C.G. Coffman, 2000a, *Cerro Grande Fire: Facility and Waste Operations Division and Facilities Lessons to be Learned Report,* Los Alamos National Laboratory Report LA-UR-01-1304.

Salazar-Langley, C.A., D.L. Hall and C.G. Coffman, 2000b, *Cerro Grande Fire: Laboratory Recovery Lessons to be Learned Report,* Los Alamos National Laboratory Report LA-UR-01-1305.

Webb, M.D. and K. Carpenter, 2001, *The Cerro Grande Fire, Los Alamos, New Mexico*, Los Alamos National Laboratory Report LA-UR-01-1630.

Witkowski, M.S., P.M. Rich, G.N. Keating, S.P. Linger, and T.L. Riggs, 2002, Spatial Data Warehouse Design for Enterprise GIS, *GIScience 2002*, Boulder, Colorado, September 25-28, 2002, 283-286.

Witkowski, M.S., P.M. Rich, and G.N. Keating, 2003, Enterprise GIS Design, in manuscript, to be submitted for peer review.